REMARKS

The Office Action mailed October 1, 2004, for the present application has been reviewed. Applicants gratefully acknowledge entry of amendments and terminal disclaimer submitted with the Request for Continued Examination filed July 15, 2004. The present amendment makes changes to the specification and to claims 1, 3, 4, 57, 59, 60, 61, 62, and 63. New claim 64 is also presented. Considered together with the following remarks, which are fully responsive to the Action, these amendments are believed sufficient to place the application into condition for allowance. No new matter has been added to the application. Applicants express appreciation for thoughtful examination by the Examiner.

Support for amendments in the specification can be found as follows:

Amendment to claim 1:	page 17, lines16-26; page 20, lines 23-27.
Amendment to claim 57:	page 5, lines 25-27; page 7, lines 22-34; page 18, lines 12-19; page 18, lines 24-34; page19, lines 10-13; page 65 (as amended) lines 5-7.
Amendment to claim 59:	Support for amendment to claim 57.
Amendment to claims 60 and 62:	page 5, lines 25-27; page 7, lines 22-34; page 18, lines 24-34.
New claim 64:	support for amendment to claim 57 and page 32, lines 30-37; page 33, lines 7-11; page 36, lines 9-12.

Sequence Compliance

The Present Action indicates that the CRF sequence listing submitted July 15, 2004, is technically flawed. An additional sequence listing was submitted September 23, 2004. A substitute sequence listing is submitted herewith, which corrects a format error identified in Raw Sequence Listing Report of September 29, 2004.

Objections to Specification.

The Present Action requests clarification concerning the submission of two different specifications. The specification as filed consists of 71 pages and corresponds to the amendments submitted July 15, 2004. As stated in the specification "as filed" at page 1, this is a Continuation-In-Part of application 08/642,712, filed May 3, 1996. A copy of the specification from the parent application, which consists of 54 pages, was filed as Exhibit A-2 in Applicants' Response to Office Action of April 9, 2003, filed October 8, 2003.

The Present Action objects to the specification as containing a Table of Contents that refers to page numbers. Applicants respectfully submit that the specification as-filed contains no Table of Contents and, accordingly, respectfully request that the objection be withdrawn.

The Action objects to use of the trademark ATCC® in the specification. Applicants have now amended the specification and respectfully request that the objection be withdrawn.

Rejections under 35 U.S.C. 112, first paragraph.

The Present Action rejects claims 1-8 and 52-54 as lacking written description in the specification. Specifically, the Action states "the claims encompass polypeptides whose molecular weight is determined by any means."

Applicants respectfully submit the claims, as amended, precisely describe the determination of molecular weight. Claims 2-8 and 52-54 read on claim 1 which, as amended, should read as follows:

1. (currently amended) An isolated OMP106 polypeptide, between 70% and 99% pure by weight, which is an outer membrane polypeptide of *Moraxella catarrhalis,* and which has a molecular weight of about 180 kD to about 230 kD as determined in SDS polyacrylamide gel electrophoresis using rabbit skeletal muscle myosin and *E. coli* β-galactosidase as the 200 kD and 116.25 kD molecular weight standards, respectively, and which comprises the amino acid sequence of SEQ ID NO: 1 or a sequence at least 80% identical to SEQ ID NO: 1.

Molecular weight is determined in SDS polyacrylamide gel electrophoresis using rabbit skeletal muscle myosin and *E. coli* β-galactosidase as the 200 kD and 116.25 kD molecular weight standards. Accordingly, Applicants respectfully request that the rejection be withdrawn.

The Action further rejects claims 57-63 as lacking written description in the specification. Specifically, the Action states "the specification fails to disclose any nucleic acid sequence that encodes an isolated polypeptide of *Moraxella catarrhalis* wherein said polypeptide is an outer membrane protein and has a molecular weight of about 180 kD to 230 kD. The teachings of the specification are limited to amino acid sequence consisting of SEQ ID NOs: 1 and 2."

Applicants respectfully submit that, in the specification as-filed, at page 31, lines 17-19, the full length OMP106 nucleic acid coding sequence and deduced amino acid sequence are disclosed. To correct errors in the sequence listing, the specification has been amended to read:

> "The nucleotide sequence of the entire OMP106 gene is depicted in SEQ ID NO: 8. A deduced amino acid sequence of the open reading frame of OMP106 is depicted in SEQ ID NO: 9."

The Action further indicates that a written description of fragments is required, above and beyond disclosure of the full length polypeptide. Citing *Vas-Cath Inc. v Mahurkar*, 19 USPQ2d 1111, the action states: "[T]he specification does not clearly allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed."

Applicants respectfully submit that the written description of fragments, as claimed, is sufficient. The "Guidelines for Examination of Patent Applications Under 35 U.S.C. 112, paragraph 1, 'Written Description' Requirement," MPEP 2163, provide [emphasis added]:

> "Possession may be shown in many ways. For example, possession may be shown ... by a clear depiction of the invention in detailed drawings or in structural chemical formulas which permit a person skilled in the art to clearly recognize that applicant had possession of the claimed invention. An adequate written description of the invention may be shown by any description of sufficient, relevant, identifying characteristics so long as a person skilled in the art would recognize that the inventor had possession

> of the claimed invention....What is conventional or well known to one of ordinary skill in the art need not be disclosed in detail...If a skilled artisan would have understood the inventor to be in possession of the claimed invention at the time of filing, even if every nuance of the claims is not explicitly described in the specification, then the adequate description requirement is met."

Applicants respectfully note that the specification discloses numerous epitopic fragments, including SEQ ID NOs: 1, 2, 12 and 13. See specification at page 24, line 36 to page 25, line 4; page 53, lines 20-23; page 56, lines 11-18. Preferred fragments, with utility as immunogens, are expressly identified as comprising epitopes of the disclosed full length OMP-106 sequence. See specification at page 19, lines 10-30; page 22, lines 22-30; page 23, lines 7-12; page 23, lines 29-30; page 24, lines 29-30; page 29, lines 34-36.

The only "fragments" claimed are those "consisting of 6 or more continuous amino acid residues" of the sequence shown in SEQ ID NO: 1 (pending claim 58). This is fully supported by the disclosure at page 24, line 36, to page 25, line 4.

Thus, Applicants respectfully submit that the claimed fragments are sufficiently described. It would be unreasonable to require amendment of the claim so as to expressly define every possible fragment of "6 or more continuous amino acid residues" of SEQ ID NO: 1.

Accordingly, Applicants respectfully request that the rejection be withdrawn.

Rejections under 35 U.S.C. 112, second paragraph.

The Action rejects claims 1-8, 52-54 and 60-63 as indefinite for failing to particularly point out and distinctly claim the subject matter.

Specifically, claims 1 and 2 are rejected as indefinite for use of the term "about" in describing a "molecular weight of about 180 kD to about 230 kD" and a "molecular weight of about 190 kD."

Applicants respectfully submit that no claim can be rejected for indefiniteness which provides a reasonable degree of particularity and distinctness. See MPEP 2173.02 [emphasis in original]. The term "about" is not indefinite *per se*. See MPEP 2173.05(b):

> "The fact that claim language, including terms of degree, may not be precise does not automatically render the claim indefinite...Acceptability of the claim language depends on whether one of ordinary skill in the art would understand what is claimed, in light of the specification...The term 'about' was held to be clear, but flexible. Ex Parte Eastwood, 163 USPQ 316."

The question is whether a person of ordinary skill in the art can interpret the metes and bounds of the claim with the term "about," so as to understand how to avoid infringement. See MPEP 2173.02.

Molecular weights determined by SDS gel electrophoresis are, by definition, approximations, referred to as "apparent" molecular weights. The determination is influenced by a variety of factors, including the size range examined. Where a broader size range is examined, distinctions become more difficult to make. Even identical gels, running identical samples under identical conditions, can produce slightly different results. Claim 1 and dependent claim 2 clearly identify two molecular weight standards of 200 kD and 116.25 kD to be used in the determination of apparent molecular weight. A person of ordinary skill in the art would clearly understand that "about" 180 kD to "about" 230 kD refers to proteins which run in such a manner as to be reasonably characterized as between 180 kD and 230 kD in a gel which clearly distinguishes between 200 kD and 116.25 kD standards. It would be unreasonable to use more definite terms to define the inherently inexact range of apparent molecular weights in SDS gels.

Accordingly, Applicants respectfully request that the rejection be withdrawn.

Claim 1, and dependent claims 5-8, 52-54, and 60-63, are further rejected as indefinite for failing to recite how molecular weight is determined. As explained above, applicants respectfully submit that claim 1 as amended, precisely describes the determination of molecular weight.

Accordingly, Applicants respectfully request that the rejection be withdrawn.

Claims 3 and 4 are rejected as indefinite for use of the trademark ATCC. Applicants have now amended these claims and respectfully request that the rejection be withdrawn.

Rejections under 35 U.S.C. 103(a)

The Action rejects all claims as obvious over Sasaki et al., US Patents Nos. 6335018, 6440424, and 6440425.

Applicants respectfully submit they have addressed the rejection and this rejection should be withdrawn, where all claims are now amended to incorporate the limitation of SEQ ID NO. 1 or its homologs. No *prima facie* obviousness can be found, where the prior art does not teach all claim limitations. See MPEP 2143.03.

SEQ ID NO. 1 is the N-terminal amino acid sequence of the mature OMP-106 polypeptide as expressed by strain ATCC 49143. See Section 6.1.7 (Isolation of OMP-106 Polypeptide) at page 46, lines 20-27 "for N-terminal sequencing;" and Section 6.2.7 (Properties of OMP106 Polypeptide) at page 53, lines 15-19, "N-terminal was sequenced." Sasaki et al. disclose a completely unrelated mature N-terminal sequence. See 6440425 Col 23, L 21-52 (SEQ ID NOs: 9 and 10). Note that Sasaki et al. report the mature polypeptide, as expressed by strain ATCC 4223, to be identical to that coded by a lambda phage expression clone from which the reported full length sequence was derived. The mature N-terminal sequence reported here, SEQ ID NO. 1, has less than 50% homology to the N-terminal sequence reported by Sasaki et al., as demonstrated by the BLAST 2 SEQUENCES alignment attached hereto as Exhibit A. [Tucker SEQ ID NO: 1 is shown as Seq 1, top; Sasaki SEQ IDs NO: 9 + 10 concatenated is shown as Seq 2, bottom].

The Action concludes that OMP-106 is an allelic variant of the Sasaki polypeptide, stating: "It would have been *prima facie* obvious to a person of ordinary skill in the art at the time the invention was made to isolate the claimed polypeptides from other strains of *Moraxella* to obtain the claimed invention."

For purposes of argument only, if Applicants accept that OMP-106 *may be* an allelic variant of the Sasaki polypeptide, this would *not* render the invention to be obvious. No *per se* rule exists that, among the genus of allelic variants, each species is "obvious." See Appeal No. 1997-3377, Ex Parte Rajender et al., Application No. 08/216326, p. 46-54 of consolidated appeals: "[T]he examiner is essentially adopting a *per se* rule that among the genus of allelic variants every species is obvious. This is clearly in error."

Even given motivation to search for allelic variants, it is *not prima facie* "obvious" that the search would lead inevitably *to the disclosed N-terminal sequence*. See *In re Deuel*, 51 F 3d 1552, 1558-1559 (Fed. Cir. 1995); and see *In re Bell*, 991 F. 2d 781, 783 (Fed. Cir. 1993). Nothing in Sasaki et al. suggests the novel N-terminal sequence disclosed here.

This novel N-terminal sequence has important functional significance. The overall sequence homology between full length OMP-106 polypeptide and the polypeptide disclosed by Sasaki et al. is only 66%, as demonstrated by the BLAST 2 SEQUENCES sequence alignment attached hereto as Exhibit B. [Tucker SEQ ID NO: 9 is shown as Seq 1, top, with the mature N-terminal sequence underlined; Sasaki SEQ ID NO: 3 is shown as Seq. 2, bottom]. As shown, while the C-terminal regions of the polypeptides have very high homology, the N-terminal regions are divergent.

In comparison with other outer membrane vaccine antigens from *Moraxella catarrhalis*, strain-to-strain variability is very high between the alleged allelic variants, the Sasaki protein and OMP-106. See e.g.:

Ref. BE US patent # 6706269 [BASB031, a cell surface adhesion antigen, has strain-to-strain sequence variability on the order of 1%; see Example 2, columns 28-30];

Ref. BF US patent #6764834 [BASB019, a cell surface adhesion antigen, has strain-to-strain sequence variability on the order of 3%; see Example 2, columns 28-30];

Ref. BG "Antigenic heterogeneity and molecular analysis of CopB of *Moraxella (Brunhamella) catarrhalis*." S. Sethl et al., Infection and Immunity (1997) 65(9):3666 [strain-to-strain sequence variability of outer membrane antigen on the order of 10%];

Ref. BH "Total genome polymorphism and low frequency of intra-genomic variation in uspA1 and uspA2 genes of *Moraxella catarrhalis* in otitis prone and non-prone children up to two years of age." J. Hays et al., Vaccine (2003), 21:1118 [low frequency of strain-to-strain sequence variability in outer membrane antigens uspA1 and upsA2].

Applicants respectfully submit the analysis of obviousness in this case does not depend on the epitopes that OMP-106 and the Sasaki protein have in common. What is significant are the dramatically *different* epitopes that *distinguish* these alleged allelic variants. Allelic variants of a rapidly evolving antigen from a rapidly evolving species are immunologically, which is to say, functionally distinct. Given the extraordinary strain-to-strain sequence variability of this antigen, OMP-106 can be compared with the Major Outer Membrane Protein (MOMP) of *Chlamydia*. The MOMP antigen is immunoprotective but genetically hypervariable. While antibodies exist that are immuno-reactive with MOMPs of all strains, antibodies against MOMP that are neutralizing, that is to say, immuno-protective, are highly strain-specific. See e.g.:

Ref. BI "Identification of protective epitopes by sequencing of the Major Outer Membrane Protein gene of a variant strain of *Chlamydia psittaci* serotype 1." E. Vretou et al., Infection and Immunity (2001) 69(1):607;

Ref. BJ "Dissociation of immune determinants of Outer Membrane Proteins of *Chlamydia psittaci* strain guinea pig inclusion conjunctivitis." T. Westbay et al., Infection and Immunity (1994) 62(12):5614;

Ref. BK "*Chlamydia trachomatis* antigens: role in immunity and pathogenesis." R. Brunham and R. Peeling, Infectious Agents and Disease (1994) 3(5):218.

The strain-to-strain sequence variability between OMP-106 and the Sasaki protein is much greater than strain-to-strain variability of *Chlamydia trachomatis* MOMP. Indeed, the homology between OMP-106 and the Sasaki protein is comparable to the homology between MOMPs from different species of *Chlamydia*. See:

Ref. BL "Comparison of the Major Outer Membrane Protein (MOMP) gene of mouse penumonitis (MoPn) and hamster SFPD strains of *Chlamydia trachomatis* with other *Chlamydia* strains." Y. Zhang et al., Mo. Biol. Evol. (1993), 10(6):1327.

Applicants respectfully note that, while OMP-106 is shown to induce production of neutralizing antibodies, Sasaki et al. present no evidence of neutralization, either *in vivo* or *in vitro*. As shown in Exhibit B, the N-terminal region of OMP-106 comprises *distinctive* epitopes. OMP-106 is functionally distinct from the Sasaki protein. Nothing in the collective works of Sasaki et al. in any way suggests the novel N-terminal epitopes reported here. Applicants respectfully submit that no *prima facie* case of obviousness has been established, even if OMP-106 is assumed to be an allelic variant of the Sasaki protein. Considered as a whole, OMP-106 as claimed is a novel and non-obvious, *Moraxella catarrhalis* neutralizing antigen.

CONCLUSION

In light of the foregoing, Applicants respectfully submit that they have addressed each and every item presented by the Examiner in this Office Action. Favorable reconsideration of all of the claims as amended is earnestly solicited. Applicants submit that the present application, with pending claims 1-8, 52-54, and 57-63 is in a condition for allowance and respectfully request such allowance.

In the event any further matters requiring attention are noted by Examiner or in the event that prosecution of this application can otherwise be advanced thereby, a telephone call to Applicants' undersigned representative at the number shown below is invited.

Respectfully submitted,

Date: January 24, 2005

Robert C. Casad, Ph.D. (Reg. No. 51,535)
Foster, Swift, Collins & Smith, P.C.
313 South Washington Square
Lansing, MI 48933
(517) 371-8100
(517) 371-8200 facsimile

Certificate of Mailing

I hereby certify that this correspondence is being deposited with the United States Postal Service Express Mail Service with sufficient postage in an envelope addressed to: , Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450.

Date: January 24, 2005

Lorri A. Rosier

Express Mail Label No. EV 458150384 US

Exhibit A

Blast 2 Sequences alignment of Tucker SEQ ID NO: 1 with SEQ ID NOs: 9 + 10 concatenated, from Sasaki et al., US Patent #6440425

Blast 2 Sequences results

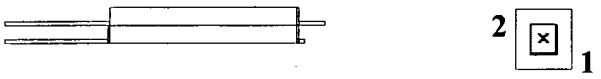
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.10 [Oct-19-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

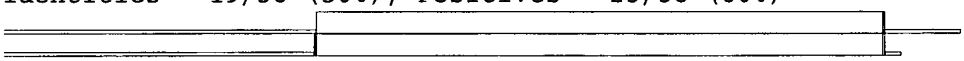
Sequence 1 lcl|seq_1 **Length** 43 (1 .. 43)

Sequence 2 lcl|seq_2 **Length** 60 (1 .. 60)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 35.0 bits (79), Expect = 0.55
Identities = 19/38 (50%), Positives = 23/38 (60%)
```



```
Query: 1  IGISEADGGKGGANARGDKSIAIGDIAQALGSQSIAIG 38
          IG        G A A GD++IAIG+ A A G Q+IAIG
Sbjct: 22 IGEQNQPRRSGTAKADGDRAIAIGENANAQGGQAIAIG 59

CPU time:     0.00 user secs.       0.02 sys. secs          0.02 total secs.

Lambda     K      H
   0.309    0.135    0.359

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 67
Number of extensions: 23
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 43
Length of database: 775,260,124
Length adjustment: 18
Effective length of query: 25
Effective length of database: 775,260,106
Effective search space: 19381502650
Effective search space used: 19381502650
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.1 bits)
X2: 129 (49.7 bits)
X3: 129 (49.7 bits)
```

S1: 42 (21.6 bits)
S2: 69 (31.2 bits)

SEQUENCES FROM TUCKER ET AL., APPLICATION 09/813214

SEQ ID NO: 1

Ile Gly Ile Ser Glu Ala Asp Gly Gly Lys Gly Gly Ala Asn Ala Arg

Gly Asp Lys Ser Ile Ala Ile Gly Asp Ile Ala Gln Ala Leu Gly Ser

Gln Ser Ile Ala Ile Gly Asp Asn Lys Ile Val

IGISEADGGKGGANARGDKS
IAIGDIAQALGSQSIAIGDN
KIV

SEQUENCES FROM SASAKI ET AL., US PATENT #6440425

SEQ ID NO: 9

Met Ile Gly Ala Thr Leu Ser Gly Ser Ala Tyr Ala Gln Lys Lys Asp

Thr Lys His Ile Ala Ile Gly Glu Gln Asn Gln Pro Arg Arg

MIGATLSGSAYAQKKDTKHI
AIGEQNQPRR

SEQ ID NO: 10

Ser Gly Thr Ala Lys Ala Asp Gly Asp Arg Ala Ile Ala Ile Gly Glu

Asn Ala Asn Ala Gln Gly Gly Gln Ala Ile Ala Ile Gly Ser

SGTAKADGDRAIAIGENANA
QGGQAIAIGS

Exhibit B

Blast 2 Sequences alignment of Tucker SEQ ID NO: 9 with SEQ ID NO: 3 from Sasaki et al., US Patent #6440425

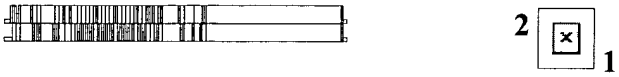
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.10 [Oct-19-2004]

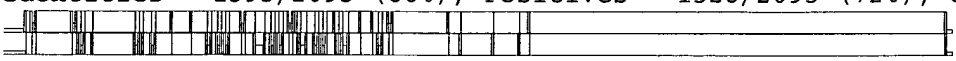
Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1 lcl|seq_1 **Length** 2122 (1 .. 2122)
Sequence 2 lcl|seq_2 **Length** 1992 (1 .. 1992)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 2434 bits (6308), Expect = 0.0
Identities = 1393/2093 (66%), Positives = 1528/2093 (72%), Gaps = 153/2093 (7%)
```



```
Query: 56   VIGATLNGSAYA--------GIGISEADGGKGGANARGDKSIAIGDIAQALGSQSIAIGD 107
            +IGATL+GSAYA          IG        G A A GD++IAIG+ A A G Q+IAIG
Sbjct: 1    MIGATLSGSAYAQKKDTKHIAIGEQNQPRRSGTAKADGDRAIAIGENANAQGGQAIAIGS 60

Query: 108  NKIVHNSNNNANIGAKASGNESIAIGGDVLASGHASIAIGSDDLYL-------KKETVQQ 160
            +    N ++    IG  A+G ESIAIGGDV ASG ASIAIGSDDL+L       K
Sbjct: 61   SNKTVNGSSLDKIGTDATGQESIAIGGDVKASGDASIAIGSDDLHLLDQHGNPKHPKGTL 120

Query: 161  ISELLPIIRGQKALNDIYQLADTNLQKYRRTHAQGHASTAVGAMSYAKGHFSNAFGTRAT 220
            I++L   I G   L +I    D ++ KYRRT A GHASTAVGAMSYA+GHFSNAFGTRAT
Sbjct: 121  INDL---INGHAVLKEIRSSKDNDV-KYRRTTASGHASTAVGAMSYAQGHFSNAFGTRAT 176

Query: 221  AEGTYSLAVGLTATAKAASSIAVGSNAQAIGFAATAVGGSTQVNLNRGIALGFGSQVLQK 280
            A+  YSLAVGL ATA+  S+IA+GS+A +    A A+G  T+  L   IALG GS V Q
Sbjct: 177  AKSAYSLAVGLAATAEGQSTIAIGSDATSSSLGAIALGAGTRAQLQGSIALGQGSVVTQS 236

Query: 281  DNDVNAANVRAYAPDDNQPIDNRYKATFKNGATDVFSIGNSNGNDSIRRKIINVGAGSAD 340
            DN+   A    Y P+  Q +D +++AT    A  +     S G++SI+RKIINVGAG
Sbjct: 237  DNNSRPA----YTPN-TQALDPKFQATNNTKAGPL-----SIGSNSIKRKIINVGAGVNK 286

Query: 341  TDAVNVAQLKEAVRLAN-RQITFKGDDSNNRVEXXXXXXXXXXXXAQTSALTDHNIGVVQ 399
            TDAVNVAQL+  V+ A  R+ITF+GDD++  V+             A+T+ALTD+NIGVV+
Sbjct: 287  TDAVNVAQLEAVVKWAKERRITFQGDDNSTDVKIGLDNTLTIKGGAETNALTDNNIGVVK 346

Query: 400  NGD--GLKVQLAETLTSLKMVTTENLTANEKVTVGKTRLTTDKIGFTNDMNGIDESKPYL 457
              D  GLKV+LA+TL +L  V T  L A   V VG +  TT ++         + +S  +
Sbjct: 347  EADNSGLKVKLAKTLNNLTEVNTTTLNATTTVKVGSSSSTTAEL--------LSDSLTFT 398

Query: 458  DKDTGIHAGGQKITKLTAGVVDDDAATYGQLKKVNQTAESALQTFTVKKVDKNG----ND 513
              +TG     Q  +K   GV          +K  N    +A    T    DK G     D
Sbjct: 399  QPNTG----SQSTSKTVYGV--------NGVKFTNNAETTAAIGTTRITRDKIGFARDGD 446

Query: 514  ANDSKIITVGKNNKPDGTQVNTLKLKGENGVDVTTETNGTVTFGLNQNNGLTVGNSTLNN 573
             ++ +   + K     G+   T+                      +NG+  GN  ++N
Sbjct: 447  VDEKQAPYLDKKQLKVGSVAITI----------------------DNGIDAGNKKISN 482

Query: 574  DGLSVKNTNSNKQIQVGADGITFTDISNSKP----GAGIENT-TRITRDGIGFANNTGSL 628
                 K +++N       D +T   +  +KP    GAGI  T T I             S+
```

```
Sbjct: 483  ---LAKGSSAN-------DAVTIEQLKAAKPTLNAGAGISVTPTEI-----------SV 520

Query: 629  DANKPRLTPTGINAGGKELTNVQSAINPATNGGQLDFMNRLSTANTEKSGSAATIKDLYN 688
            DA    +T    N G K             + G  D   + S   +  + S  T + L +
Sbjct: 521  DAKSGNVTAPTYNIGVKT--------TELNSDGTSD---KFSVKGSGTNNSLVTAEHLAS 569

Query: 689  -LSQVPLTFAGDTGPNVTKKLGEILKVKGGKTTADDLTKNNIGVVADSTDNSLTVKLAKT 747
             L++V  T A     + T K               DD   N I V  D+T N+  V + K
Sbjct: 570  YLNEVNRT-ADSALQSFTVK-----------EEDDDDANAITVAKDTTKNAGAVSILKL 616

Query: 748  L--SDLDAVNTKTLTASDKVTVDSGNNTAK--LQNGDLTFSKQNTGATPATNSKTIGVDG 803
               + L     K  T +  ++ DSG    K  L N  LT    N           +G +G
Sbjct: 617  KGKNGLTVATKKDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDTN-------EQIQVGANG 669

Query: 804  LKFTD----NNGIALDGTTYITKDKVGFAKQDGSLDKSKPYLDKDKLKVGEVEITTNGIN 859
            +KFT+    N G  +  T  IT+DK+GFA  DG++D +KPYLD+DKL+VG V+IT  GIN
Sbjct: 670  IKFTNVNGSNPGTGIANTARITRDKIGFAGSDGAVDTNKPYLDQDKLQVGNVKITNTGIN 729

Query: 860  AGGKAITGLSNTLTDATNATTGHVTQLGIVDSTDKTRAASIGDVLNAGFNLKNNGDAKDF 919
            AGGKAITGLS TL    + ++ ++     +   DK+ AASI D+LN GFNLKNN +  DF
Sbjct: 730  AGGKAITGLSPTLPSIADQSSRNIELGNTIQDKDKSNAASINDILNTGFNLKNNNNPIDF 789

Query: 920  VSTYDTVDFINGNATTAKVTYD--GKASKVAYDVNVDGTTIHLTGADGNKNQIGVKTTTL 977
            VSTYD VDF NGNATTA VT+D   K SKV YDVNVD TTIHLTG D NK ++GVKTT L
Sbjct: 790  VSTYDIVDFANGNATTATVTHDTANKTSKVVYDVNVDDTTIHLTGTDDNK-KLGVKTTKL 848

Query: 978  TKTDAKGDKAINFSVNSGDDKALINAKDIADNLNTLAGEIRNTKGTADTALQTFQVKKVK 1037
             KT A G+ A NF+VNS D+ AL+NAKDIA+NLNTLA EI  TKGTADTALQTF VKKV
Sbjct: 849  NKTSANGNTATNFNVNSSDEDALVNAKDIAENLNTLAKEIHTTKGTADTALQTFTVKKVD 908

Query: 1038 ENGDDDNDADTITVGKDAKTNQVNTLKLKGKNGLDIQTNKDGTVTFGINTQSGLKAXXXX 1097
            EN + D DA+ ITVG+    NQVNTL LKG+NGL+I+T+K+GTVTFGINT SGLKA
Sbjct: 909  ENNNAD-DANAITVGQKNANNQVNTLTLKGENGLNIKTDKNGTVTFGINTTSGLKAGKST 967

Query: 1098 XXXXXXXSIKNTAGNEQIQVGADGVKFAKVNN-GVVGAGIDGTTRITRDEIGFAGTNGSL 1156
                  SIKN  G+EQIQVGADGVKFAKVNN GVVGAGIDGTTRITRDEIGF GTNGSL
Sbjct: 968  LNDGGL-SIKNPTGSEQIQVGADGVKFAKVNNNGVVGAGIDGTTRITRDEIGFTGTNGSL 1026

Query: 1157 DKSKPHLSKDGINAGGKKITNIQSGEIAQNSNDAVTGGKIYDLKTELENKISSTAKTAQN 1216
            DKSKPHLSKDGINAGGKKITNIQSGEIAQNS+DAVTGGKIYDLKTELENKISSTAKTAQN
Sbjct: 1027 DKSKPHLSKDGINAGGKKITNIQSGEIAQNSHDAVTGGKIYDLKTELENKISSTAKTAQN 1086

Query: 1217 SLHEFSVADEQGNNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNKGVVRVGIDQTKGL 1276
            SLHEFSVADEQGNNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNKGVVRVGIDQTKGL
Sbjct: 1087 SLHEFSVADEQGNNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNKGVVRVGIDQTKGL 1146

Query: 1277 TTPKLTVGNNNGKGIVIDSQNGQNTITGLSNTLANVTNDKGSVRTTEQGKIIKDEDKTRA 1336
            TTPKLTVGNNNGKGIVIDSQNGQNTITGLSNTLANVTNDKGSVRTTEQG IIKDEDKTRA
Sbjct: 1147 TTPKLTVGNNNGKGIVIDSQNGQNTITGLSNTLANVTNDKGSVRTTEQGNIIKDEDKTRA 1206

Query: 1337 ASIVDVLSAGFNLQGNGEAVDFVSTYDTVNFADGNATTAKVTYDDTSKTSKVVYDVNVDD 1396
            ASIVDVLSAGFNLQGNGEAVDFVSTYDTVNFADGNATTAKVTYDDTSKTSKVVYDVNVDD
Sbjct: 1207 ASIVDVLSAGFNLQGNGEAVDFVSTYDTVNFADGNATTAKVTYDDTSKTSKVVYDVNVDD 1266

Query: 1397 TTIEVKDXXXXXXXXXXXXXXXXANKFALSNQATGDALVKASDIVAHLNTLSGDIQTAKG 1456
            TTIEVKD                ANKFALSNQATGDALVKASDIVAHLNTLSGDIQTAKG
Sbjct: 1267 TTIEVKDKKLGVKTTTLTSTGTGANKFALSNQATGDALVKASDIVAHLNTLSGDIQTAKG 1326

Query: 1457 ASQANSSAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTKEVAKDKLVAQAQTPDGTLAQ 1516
            ASQAN+SAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTKEVAKDKLVAQAQTPDGTLAQ
Sbjct: 1327 ASQANNSAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTKEVAKDKLVAQAQTPDGTLAQ 1386

Query: 1517 MNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAASDNKTKNAAVTVGDLNAVAQTPLTF 1576
            MNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAASDNKTKNAAVTVGDLNAVAQTPLTF
```

```
Sbjct: 1387 MNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAASDNKTKNAAVTVGDLNAVAQTPLTF 1446

Query: 1577 AGDTGTTAKKLGETLTIKGGQTDTNKLTDNNIGVVAGTDGFTVKLAKDLTNLNSVNAGGT 1636
            AGDTGTTAKKLGETLTIKGGQTDTNKLTDNNIGVVAGTDGFTVKLAKDLTNLNSVNAGGT
Sbjct: 1447 AGDTGTTAKKLGETLTIKGGQTDTNKLTDNNIGVVAGTDGFTVKLAKDLTNLNSVNAGGT 1506

Query: 1637 KIDDKGVSFVDSSGQAKANTPVLSANGLDLGGKVISNVGKGTKDTDAANVQQLNEVRXXX 1696
            KIDDKGVSFVDSSGQAKANTPVLSANGLDLGGKVISNVGKGTKDTDAANVQQLNEVR
Sbjct: 1507 KIDDKGVSFVDSSGQAKANTPVLSANGLDLGGKVISNVGKGTKDTDAANVQQLNEVRNLL 1566

Query: 1697 XXXXXXXXXXXXXQVNIADIKKDPNSGSSSNRTVIKAGTVLGGKGNNDTEKLATGGVQVG 1756
                          QVNIADIKKDPNSGSSSNRTVIKAGTVLGGKGNNDTEKLATGG+QVG
Sbjct: 1567 GLGNAGNDNADGNQVNIADIKKDPNSGSSSNRTVIKAGTVLGGKGNNDTEKLATGGIQVG 1626

Query: 1757 VDKDGNANGDLSNVWVKTQKDGSKKALLATYNAAGQTNYLTNNPAEAIDRINEQGIRFFH 1816
            VDKDGNANGDLSNVWVKTQKDGSKKALLATYNAAGQTNYLTNNPAEAIDRINEQGIRFFH
Sbjct: 1627 VDKDGNANGDLSNVWVKTQKDGSKKALLATYNAAGQTNYLTNNPAEAIDRINEQGIRFFH 1686

Query: 1817 VNDGNQEPVVQGRNGIDSSASGKHSVAIGFQAKADGEAAVAIGRQTQAGNQSIAIGDNAQ 1876
            VNDGNQEPVVQGRNGIDSSASGKHSVAIGFQAKADGEAAVAIGRQTQAGNQSIAIGDNAQ
Sbjct: 1687 VNDGNQEPVVQGRNGIDSSASGKHSVAIGFQAKADGEAAVAIGRQTQAGNQSIAIGDNAQ 1746

Query: 1877 ATGDQSIAIGTGNVVAGKHSGAIGDPSTVKADNSYSVGNNNQFTDATQTDVFGVGNNITV 1936
            ATGDQSIAIGTGNVVAGKHSGAIGDPSTVKADNSYSVGNNNQFTDATQTDVFGVGNNITV
Sbjct: 1747 ATGDQSIAIGTGNVVAGKHSGAIGDPSTVKADNSYSVGNNNQFTDATQTDVFGVGNNITV 1806

Query: 1937 TESNSVALGSNSAISAGTHAGTQAKKSDXXXXXXXXXXXXXXVKGFAGQTXXXXXXXXXX 1996
            TESNSVALGSNSAISAGTHAGTQAKKSD               VKGFAGQT
Sbjct: 1807 TESNSVALGSNSAISAGTHAGTQAKKSDGTAGTTTTAGATGTVKGFAGQTAVGAVSVGAS 1866

Query: 1997 XXERRIQNVAAGEVSATSTDAVNGSQLYKATQGIANATNELDHRIHQNENKANAGIXXXX 2056
              ERRIQNVAAGEVSATSTDAVNGSQLYKATQ IANATNELDHRIHQNENKANAGI
Sbjct: 1867 GAERRIQNVAAGEVSATSTDAVNGSQLYKATQSIANATNELDHRIHQNENKANAGISSAM 1926

Query: 2057 XXXXXPQAYIPGRSMVTGGIATHNGQGAVAVGLSKLSDNGQWVFKINGSADTQ 2109
                 PQAYIPGRSMVTGGIATHNGQGAVAVGLSKLSDNGQWVFKINGSADTQ
Sbjct: 1927 AMASMPQAYIPGRSMVTGGIATHNGQGAVAVGLSKLSDNGQWVFKINGSADTQ 1979

CPU time:     0.24 user secs.       0.01 sys. secs          0.25 total secs.

Lambda     K      H
   0.307    0.126    0.339

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 48,642
Number of extensions: 28435
Number of successful extensions: 179
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 5
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 2122
Length of database: 775,260,124
Length adjustment: 146
Effective length of query: 1976
Effective length of database: 775,259,978
```

```
Effective search space: 1531913716528
Effective search space used: 1531913716528
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.1 bits)
X2: 129 (49.7 bits)
X3: 129 (49.7 bits)
S1: 42 (21.6 bits)
S2: 85 (37.4 bits)
```

SEQUENCES FROM TUCKER ET AL., APPLICATION 09/813214

SEQ ID NO: 9 [as amended]

Met Asn His Ile Tyr Lys Val Ile Phe Asn Lys Ala Thr Gly Thr Phe

Met Ala Val Ala Glu Tyr Ala Lys Ser His Ser Thr Gly Gly Gly Ser

Cys Ala Thr Gly Gln Val Gly Ser Val Arg Thr Leu Ser Phe Ala Arg

Ile Ala Ala Leu Ala Val Leu Val Ile Gly Ala Thr Leu Asn Gly Ser

Ala Tyr Ala Gly Ile Gly Ile Ser Glu Ala Asp Gly Gly Lys Gly Gly

Ala Asn Ala Arg Gly Asp Lys Ser Ile Ala Ile Gly Asp Ile Ala Gln

Ala Leu Gly Ser Gln Ser Ile Ala Ile Gly Asp Asn Lys Ile Val His

Asn Ser Asn Asn Asn Ala Asn Ile Gly Ala Lys Ala Ser Gly Asn Glu

Ser Ile Ala Ile Gly Gly Asp Val Leu Ala Ser Gly His Ala Ser Ile

Ala Ile Gly Ser Asp Asp Leu Tyr Leu Lys Lys Glu Thr Val Gln Gln

Ile Ser Glu Leu Leu Pro Ile Ile Arg Gly Gln Lys Ala Leu Asn Asp

Ile Tyr Gln Leu Ala Asp Thr Asn Leu Gln Lys Tyr Arg Arg Thr His

Ala Gln Gly His Ala Ser Thr Ala Val Gly Ala Met Ser Tyr Ala Lys

Gly His Phe Ser Asn Ala Phe Gly Thr Arg Ala Thr Ala Glu Gly Thr

Tyr Ser Leu Ala Val Gly Leu Thr Ala Thr Ala Lys Ala Ala Ser Ser

Ile Ala Val Gly Ser Asn Ala Gln Ala Ile Gly Phe Ala Ala Thr Ala

Val Gly Gly Ser Thr Gln Val Asn Leu Asn Arg Gly Ile Ala Leu Gly

Phe Gly Ser Gln Val Leu Gln Lys Asp Asn Asp Val Asn Ala Ala Asn

Val Arg Ala Tyr Ala Pro Asp Asp Asn Gln Pro Ile Asp Asn Arg Tyr

Lys Ala Thr Phe Lys Asn Gly Ala Thr Asp Val Phe Ser Ile Gly Asn

Ser Asn Gly Asn Asp Ser Ile Arg Arg Lys Ile Ile Asn Val Gly Ala

Gly Ser Ala Asp Thr Asp Ala Val Asn Val Ala Gln Leu Lys Glu Ala

Val Arg Leu Ala Asn Arg Gln Ile Thr Phe Lys Gly Asp Asp Ser Asn

Asn Arg Val Glu Lys Gly Leu Gly Lys Thr Leu Thr Ile Thr Gly Gly

Ala Gln Thr Ser Ala Leu Thr Asp His Asn Ile Gly Val Val Gln Asn

Gly Asp Gly Leu Lys Val Gln Leu Ala Glu Thr Leu Thr Ser Leu Lys

Met Val Thr Thr Glu Asn Leu Thr Ala Asn Glu Lys Val Thr Val Gly

Lys Thr Arg Leu Thr Thr Asp Lys Ile Gly Phe Thr Asn Asp Met Asn

Gly Ile Asp Glu Ser Lys Pro Tyr Leu Asp Lys Asp Thr Gly Ile His

Ala Gly Gly Gln Lys Ile Thr Lys Leu Thr Ala Gly Val Val Asp Asp

Asp Ala Ala Thr Tyr Gly Gln Leu Lys Lys Val Asn Gln Thr Ala Glu

Ser Ala Leu Gln Thr Phe Thr Val Lys Lys Val Asp Lys Asn Gly Asn

Asp Ala Asn Asp Ser Lys Ile Ile Thr Val Gly Lys Asn Asn Lys Pro

Asp Gly Thr Gln Val Asn Thr Leu Lys Leu Lys Gly Glu Asn Gly Val

Asp Val Thr Thr Glu Thr Asn Gly Thr Val Thr Phe Gly Leu Asn Gln

Asn Asn Gly Leu Thr Val Gly Asn Ser Thr Leu Asn Asn Asp Gly Leu

Ser Val Lys Asn Thr Asn Ser Asn Lys Gln Ile Gln Val Gly Ala Asp

Gly Ile Thr Phe Thr Asp Ile Ser Asn Ser Lys Pro Gly Ala Gly Ile

Glu Asn Thr Thr Arg Ile Thr Arg Asp Gly Ile Gly Phe Ala Asn Asn

Thr Gly Ser Leu Asp Ala Asn Lys Pro Arg Leu Thr Pro Thr Gly Ile

Asn Ala Gly Gly Lys Glu Leu Thr Asn Val Gln Ser Ala Ile Asn Pro

Ala Thr Asn Gly Gly Gln Leu Asp Phe Met Asn Arg Leu Ser Thr Ala

Asn Thr Glu Lys Ser Gly Ser Ala Ala Thr Ile Lys Asp Leu Tyr Asn

Leu Ser Gln Val Pro Leu Thr Phe Ala Gly Asp Thr Gly Pro Asn Val

Thr Lys Lys Leu Gly Glu Ile Leu Lys Val Lys Gly Gly Lys Thr Thr

Ala Asp Asp Leu Thr Lys Asn Asn Ile Gly Val Val Ala Asp Ser Thr

Asp Asn Ser Leu Thr Val Lys Leu Ala Lys Thr Leu Ser Asp Leu Asp

Ala Val Asn Thr Lys Thr Leu Thr Ala Ser Asp Lys Val Thr Val Asp

Ser Gly Asn Asn Thr Ala Lys Leu Gln Asn Gly Asp Leu Thr Phe Ser

Lys Gln Asn Thr Gly Ala Thr Pro Ala Thr Asn Ser Lys Thr Ile Gly

Val Asp Gly Leu Lys Phe Thr Asp Asn Asn Gly Ile Ala Leu Asp Gly

Thr Thr Tyr Ile Thr Lys Asp Lys Val Gly Phe Ala Lys Gln Asp Gly

Ser Leu Asp Lys Ser Lys Pro Tyr Leu Asp Lys Asp Lys Leu Lys Val

Gly Glu Val Glu Ile Thr Thr Asn Gly Ile Asn Ala Gly Gly Lys Ala

Ile Thr Gly Leu Ser Asn Thr Leu Thr Asp Ala Thr Asn Ala Thr Thr

Gly His Val Thr Gln Leu Gly Ile Val Asp Ser Thr Asp Lys Thr Arg

Ala Ala Ser Ile Gly Asp Val Leu Asn Ala Gly Phe Asn Leu Lys Asn

Asn Gly Asp Ala Lys Asp Phe Val Ser Thr Tyr Asp Thr Val Asp Phe

Ile Asn Gly Asn Ala Thr Thr Ala Lys Val Thr Tyr Asp Gly Lys Ala

Ser Lys Val Ala Tyr Asp Val Asn Val Asp Gly Thr Thr Ile His Leu

Thr Gly Ala Asp Gly Asn Lys Asn Gln Ile Gly Val Lys Thr Thr Thr

Leu Thr Lys Thr Asp Ala Lys Gly Asp Lys Ala Ile Asn Phe Ser Val

Asn Ser Gly Asp Asp Lys Ala Leu Ile Asn Ala Lys Asp Ile Ala Asp

Asn Leu Asn Thr Leu Ala Gly Glu Ile Arg Asn Thr Lys Gly Thr

Ala Asp Thr Ala Leu Gln Thr Phe Gln Val Lys Lys Val Lys Glu

Asn Gly Asp Asp Asp Asn Asp Ala Asp Thr Ile Thr Val Gly Lys

Asp Ala Lys Thr Asn Gln Val Asn Thr Leu Lys Leu Lys Gly Lys

Asn Gly Leu Asp Ile Gln Thr Asn Lys Asp Gly Thr Val Thr Phe

Gly Ile Asn Thr Gln Ser Gly Leu Lys Ala Gly Asn Asn Thr Thr

Leu Asn Asn Asn Gly Leu Ser Ile Lys Asn Thr Ala Gly Asn Glu

Gln Ile Gln Val Gly Ala Asp Gly Val Lys Phe Ala Lys Val Asn

Asn Gly Val Val Gly Ala Gly Ile Asp Gly Thr Thr Arg Ile Thr

Arg Asp Glu Ile Gly Phe Ala Gly Thr Asn Gly Ser Leu Asp Lys

Ser Lys Pro His Leu Ser Lys Asp Gly Ile Asn Ala Gly Gly Lys

Lys Ile Thr Asn Ile Gln Ser Gly Glu Ile Ala Gln Asn Ser Asn

Asp Ala Val Thr Gly Gly Lys Ile Tyr Asp Leu Lys Thr Glu Leu

Glu Asn Lys Ile Ser Ser Thr Ala Lys Thr Ala Gln Asn Ser Leu

His Glu Phe Ser Val Ala Asp Glu Gln Gly Asn Asn Phe Thr Val

Ser Asn Pro Tyr Ser Ser Tyr Asp Thr Ser Lys Thr Ser Asp Val

Ile Thr Phe Ala Gly Glu Asn Gly Ile Thr Thr Lys Val Asn Lys

Gly Val Val Arg Val Gly Ile Asp Gln Thr Lys Gly Leu Thr Thr

Pro Lys Leu Thr Val Gly Asn Asn Asn Gly Lys Gly Ile Val Ile

Asp Ser Gln Asn Gly Gln Asn Thr Ile Thr Gly Leu Ser Asn Thr

Leu Ala Asn Val Thr Asn Asp Lys Gly Ser Val Arg Thr Thr Glu

Gln Gly Lys Ile Ile Lys Asp Glu Asp Lys Thr Arg Ala Ala Ser

Ile Val Asp Val Leu Ser Ala Gly Phe Asn Leu Gln Gly Asn Gly

Glu Ala Val Asp Phe Val Ser Thr Tyr Asp Thr Val Asn Phe Ala

Asp Gly Asn Ala Thr Thr Ala Lys Val Thr Tyr Asp Asp Thr Ser

Lys Thr Ser Lys Val Val Tyr Asp Val Asn Val Asp Asp Thr Thr

Ile Glu Val Lys Asp Lys Lys Leu Gly Val Lys Thr Thr Thr Leu

Thr Ser Thr Gly Thr Gly Ala Asn Lys Phe Ala Leu Ser Asn Gln

Ala Thr Gly Asp Ala Leu Val Lys Ala Ser Asp Ile Val Ala His

Leu Asn Thr Leu Ser Gly Asp Ile Gln Thr Ala Lys Gly Ala Ser

Gln Ala Asn Ser Ser Ala Gly Tyr Val Asp Ala Asp Gly Asn Lys

Val Ile Tyr Asp Ser Thr Asp Asn Lys Tyr Tyr Gln Ala Lys Asn

Asp Gly Thr Val Asp Lys Thr Lys Glu Val Ala Lys Asp Lys Leu

Val Ala Gln Ala Gln Thr Pro Asp Gly Thr Leu Ala Gln Met Asn

Val Lys Ser Val Ile Asn Lys Glu Gln Val Asn Asp Ala Asn Lys

Lys Gln Gly Ile Asn Glu Asp Asn Ala Phe Val Lys Gly Leu Glu

Lys Ala Ala Ser Asp Asn Lys Thr Lys Asn Ala Ala Val Thr Val

Gly Asp Leu Asn Ala Val Ala Gln Thr Pro Leu Thr Phe Ala Gly

Asp Thr Gly Thr Thr Ala Lys Lys Leu Gly Glu Thr Leu Thr Ile

Lys Gly Gly Gln Thr Asp Thr Asn Lys Leu Thr Asp Asn Asn Ile

Gly Val Val Ala Gly Thr Asp Gly Phe Thr Val Lys Leu Ala Lys

Asp Leu Thr Asn Leu Asn Ser Val Asn Ala Gly Gly Thr Lys Ile

Asp Asp Lys Gly Val Ser Phe Val Asp Ser Ser Gly Gln Ala Lys

Ala Asn Thr Pro Val Leu Ser Ala Asn Gly Leu Asp Leu Gly Gly

Lys Val Ile Ser Asn Val Gly Lys Gly Thr Lys Asp Thr Asp Ala

Ala Asn Val Gln Gln Leu Asn Glu Val Arg Asn Leu Leu Gly Leu

Gly Asn Ala Gly Asn Asp Asn Ala Asp Gly Asn Gln Val Asn Ile

Ala Asp Ile Lys Lys Asp Pro Asn Ser Gly Ser Ser Ser Asn Arg

Thr Val Ile Lys Ala Gly Thr Val Leu Gly Gly Lys Gly Asn Asn

Asp Thr Glu Lys Leu Ala Thr Gly Gly Val Gln Val Gly Val Asp

Lys Asp Gly Asn Ala Asn Gly Asp Leu Ser Asn Val Trp Val Lys

Thr Gln Lys Asp Gly Ser Lys Lys Ala Leu Leu Ala Thr Tyr Asn

Ala Ala Gly Gln Thr Asn Tyr Leu Thr Asn Asn Pro Ala Glu Ala

Ile Asp Arg Ile Asn Glu Gln Gly Ile Arg Phe Phe His Val Asn

Asp Gly Asn Gln Glu Pro Val Val Gln Gly Arg Asn Gly Ile Asp

Ser Ser Ala Ser Gly Lys His Ser Val Ala Ile Gly Phe Gln Ala

Lys Ala Asp Gly Glu Ala Ala Val Ala Ile Gly Arg Gln Thr Gln

Ala Gly Asn Gln Ser Ile Ala Ile Gly Asp Asn Ala Gln Ala Thr

Gly Asp Gln Ser Ile Ala Ile Gly Thr Gly Asn Val Val Ala Gly

Lys His Ser Gly Ala Ile Gly Asp Pro Ser Thr Val Lys Ala Asp

Asn Ser Tyr Ser Val Gly Asn Asn Asn Gln Phe Thr Asp Ala Thr

Gln Thr Asp Val Phe Gly Val Gly Asn Asn Ile Thr Val Thr Glu

Ser Asn Ser Val Ala Leu Gly Ser Asn Ser Ala Ile Ser Ala Gly

Thr His Ala Gly Thr Gln Ala Lys Lys Ser Asp Gly Thr Ala Gly

Thr Thr Thr Thr Ala Gly Ala Thr Gly Thr Val Lys Gly Phe Ala

Gly Gln Thr Ala Val Gly Ala Val Ser Val Gly Ala Ser Gly Ala

Glu Arg Arg Ile Gln Asn Val Ala Ala Gly Glu Val Ser Ala Thr

Ser Thr Asp Ala Val Asn Gly Ser Gln Leu Tyr Lys Ala Thr Gln

Gly Ile Ala Asn Ala Thr Asn Glu Leu Asp His Arg Ile His Gln

Asn Glu Asn Lys Ala Asn Ala Gly Ile Ser Ser Ala Met Ala Met

Ala Ser Met Pro Gln Ala Tyr Ile Pro Gly Arg Ser Met Val Thr

Gly Gly Ile Ala Thr His Asn Gly Gln Gly Ala Val Ala Val Gly

Leu Ser Lys Leu Ser Asp Asn Gly Gln Trp Val Phe Lys Ile Asn

Gly Ser Ala Asp Thr Gln Gly His Val Gly Ala Ala Val Gly Ala

Gly Phe His Phe

MNHIYKVIFNKATGTFMAVA
EYAKSHSTGGGSCATGQVGS
VRTLSFARIAALAVLVIGAT
LNGSAYAGIGISEADGGKGG
ANARGDKSIAIGDIAQALGS
QSIAIGDNKIVHNSNNNANI
GAKASGNESIAIGGDVLASG
HASIAIGSDDLYLKKETVQQ
ISELLPIIRGQKALNDIYQL
ADTNLQKYRRTHAQGHASTA
VGAMSYAKGHFSNAFGTRAT
AEGTYSLAVGLTATAKAASS
IAVGSNAQAIGFAATAVGGS
TQVNLNRGIALGFGSQVLQK
DNDVNAANVRAYAPDDNQPI
DNRYKATFKNGATDVFSIGN
SNGNDSIRRKIINVGAGSAD
TDAVNVAQLKEAVRLANRQI
TFKGDDSNNRVEKGLGKTLT
ITGGAQTSALTDHNIGVVQN
GDGLKVQLAETLTSLKMVTT
ENLTANEKVTVGKTRLTTDK
IGFTNDMNGIDESKPYLDKD
TGIHAGGQKITKLTAGVVDD
DAATYGQLKKVNQTAESALQ
TFTVKKVDKNGNDANDSKII
TVGKNNKPDGTQVNTLKLKG
ENGVDVTTETNGTVTFGLNQ
NNGLTVGNSTLNNDGLSVKN
TNSNKQIQVGADGITFTDIS
NSKPGAGIENTTRITRDGIG
FANNTGSLDANKPRLTPTGI
NAGGKELTNVQSAINPATNG
GQLDFMNRLSTANTEKSGSA
ATIKDLYNLSQVPLTFAGDT
GPNVTKKLGEILKVKGGKTT
ADDLTKNNIGVVADSTDNSL
TVKLAKTLSDLDAVNTKTLT
ASDKVTVDSGNNTAKLQNGD
LTFSKQNTGATPATNSKTIG
VDGLKFTDNNGIALDGTTYI
TKDKVGFAKQDGSLDKSKPY
LDKDKLKVGEVEITTNGINA
GGKAITGLSNTLTDATNATT
GHVTQLGIVDSTDKTRAASI
GDVLNAGFNLKNNGDAKDFV
STYDTVDFINGNATTAKVTY
DGKASKVAYDVNVDGTTIHL
TGADGNKNQIGVKTTTLTKT
DAKGDKAINFSVNSGDDKAL
INAKDIADNLNTLAGEIRNT
KGTADTALQTFQVKKVKENG
DDDNDADTITVGKDAKTNQV

NTLKLKGKNGLDIQTNKDGT
VTFGINTQSGLKAGNNTTLN
NNGLSIKNTAGNEQIQVGAD
GVKFAKVNNGVVGAGIDGTT
RITRDEIGFAGTNGSLDKSK
PHLSKDGINAGGKKITNIQS
GEIAQNSNDAVTGGKIYDLK
TELENKISSTAKTAQNSLHE
FSVADEQGNNFTVSNPYSSY
DTSKTSDVITFAGENGITTK
VNKGVVRVGIDQTKGLTTPK
LTVGNNNGKGIVIDSQNGQN
TITGLSNTLANVTNDKGSVR
TTEQGKIIKDEDKTRAASIV
DVLSAGFNLQGNGEAVDFVS
TYDTVNFADGNATTAKVTYD
DTSKTSKVVYDVNVDDTTIE
VKDKKLGVKTTTLTSTGTGA
NKFALSNQATGDALVKASDI
VAHLNTLSGDIQTAKGASQA
NSSAGYVDADGNKVIYDSTD
NKYYQAKNDGTVDKTKEVAK
DKLVAQAQTPDGTLAQMNVK
SVINKEQVNDANKKQGINED
NAFVKGLEKAASDNKTKNAA
VTVGDLNAVAQTPLTFAGDT
GTTAKKLGETLTIKGGQTDT
NKLTDNNIGVVAGTDGFTVK
LAKDLTNLNSVNAGGTKIDD
KGVSFVDSSGQAKANTPVLS
ANGLDLGGKVISNVGKGTKD
TDAANVQQLNEVRNLLGLGN
AGNDNADGNQVNIADIKKDP
NSGSSSNRTVIKAGTVLGGK
GNNDTEKLATGGVQVGVDKD
GNANGDLSNVWVKTQKDGSK
KALLATYNAAGQTNYLTNNP
AEAIDRINEQGIRFFHVNDG
NQEPVVQGRNGIDSSASGKH
SVAIGFQAKADGEAAVAIGR
QTQAGNQSIAIGDNAQATGD
QSIAIGTGNVVAGKHSGAIG
DPSTVKADNSYSVGNNNQFT
DATQTDVFGVGNNITVTESN
SVALGSNSAISAGTHAGTQA
KKSDGTAGTTTTAGATGTVK
GFAGQTAVGAVSVGASGAER
RIQNVAAGEVSATSTDAVNG
SQLYKATQGIANATNELDHR
IHQNENKANAGISSAMAMAS
MPQAYIPGRSMVTGGIATHN
GQGAVAVGLSKLSDNGQWVF
KINGSADTQGHVGAAVGAGF
HF

SEQUENCES FROM SASAKI ET AL., US PATENT #6440425

SEQ ID NO: 3

Met Ile Gly Ala Thr Leu Ser Gly Ser Ala Tyr Ala Gln Lys Lys Asp
Thr Lys His Ile Ala Ile Gly Glu Gln Asn Gln Pro Arg Arg Ser Gly
Thr Ala Lys Ala Asp Gly Asp Arg Ala Ile Ala Ile Gly Glu Asn Ala
Asn Ala Gln Gly Gly Gln Ala Ile Ala Ile Gly Ser Ser Asn Lys Thr
Val Asn Gly Ser Ser Leu Asp Lys Ile Gly Thr Asp Ala Thr Gly Gln
Glu Ser Ile Ala Ile Gly Gly Asp Val Lys Ala Ser Gly Asp Ala Ser
Ile Ala Ile Gly Ser Asp Asp Leu His Leu Leu Asp Gln His Gly Asn
Pro Lys His Pro Lys Gly Thr Leu Ile Asn Asp Leu Ile Asn Gly His
Ala Val Leu Lys Glu Ile Arg Ser Ser Lys Asp Asn Asp Val Lys Tyr
Arg Arg Thr Thr Ala Ser Gly His Ala Ser Thr Ala Val Gly Ala Met
Ser Tyr Ala Gln Gly His Phe Ser Asn Ala Phe Gly Thr Arg Ala Thr
Ala Lys Ser Ala Tyr Ser Leu Ala Val Gly Leu Ala Ala Thr Ala Glu
Gly Gln Ser Thr Ile Ala Ile Gly Ser Asp Ala Thr Ser Ser Ser Leu
Gly Ala Ile Ala Leu Gly Ala Gly Thr Arg Ala Gln Leu Gln Gly Ser
Ile Ala Leu Gly Gln Gly Ser Val Val Thr Gln Ser Asp Asn Asn Ser
Arg Pro Ala Tyr Thr Pro Asn Thr Gln Ala Leu Asp Pro Lys Phe Gln
Ala Thr Asn Asn Thr Lys Ala Gly Pro Leu Ser Ile Gly Ser Asn Ser
Ile Lys Arg Lys Ile Ile Asn Val Gly Ala Gly Val Asn Lys Thr Asp
Ala Val Asn Val Ala Gln Leu Glu Ala Val Val Lys Trp Ala Lys Glu
Arg Arg Ile Thr Phe Gln Gly Asp Asp Asn Ser Thr Asp Val Lys Ile
Gly Leu Asp Asn Thr Leu Thr Ile Lys Gly Gly Ala Glu Thr Asn Ala
Leu Thr Asp Asn Asn Ile Gly Val Val Lys Glu Ala Asp Asn Ser Gly
Leu Lys Val Lys Leu Ala Lys Thr Leu Asn Asn Leu Thr Glu Val Asn
Thr Thr Thr Leu Asn Ala Thr Thr Thr Val Lys Val Gly Ser Ser Ser
Ser Thr Thr Ala Glu Leu Leu Ser Asp Ser Leu Thr Phe Thr Gln Pro
Asn Thr Gly Ser Gln Ser Thr Ser Lys Thr Val Tyr Gly Val Asn Gly

Val Lys Phe Thr Asn Asn Ala Glu Thr Thr Ala Ala Ile Gly Thr Thr

Arg Ile Thr Arg Asp Lys Ile Gly Phe Ala Arg Asp Gly Asp Val Asp

Glu Lys Gln Ala Pro Tyr Leu Asp Lys Lys Gln Leu Lys Val Gly Ser

Val Ala Ile Thr Ile Asp Asn Gly Ile Asp Ala Gly Asn Lys Lys Ile

Ser Asn Leu Ala Lys Gly Ser Ser Ala Asn Asp Ala Val Thr Ile Glu

Gln Leu Lys Ala Ala Lys Pro Thr Leu Asn Ala Gly Ala Gly Ile Ser

Val Thr Pro Thr Glu Ile Ser Val Asp Ala Lys Ser Gly Asn Val Thr

Ala Pro Thr Tyr Asn Ile Gly Val Lys Thr Thr Glu Leu Asn Ser Asp

Gly Thr Ser Asp Lys Phe Ser Val Lys Gly Ser Gly Thr Asn Asn Ser

Leu Val Thr Ala Glu His Leu Ala Ser Tyr Leu Asn Glu Val Asn Arg

Thr Ala Asp Ser Ala Leu Gln Ser Phe Thr Val Lys Glu Glu Asp Asp

Asp Asp Ala Asn Ala Ile Thr Val Ala Lys Asp Thr Thr Lys Asn Ala

Gly Ala Val Ser Ile Leu Lys Leu Lys Gly Lys Asn Gly Leu Thr Val

Ala Thr Lys Lys Asp Gly Thr Val Thr Phe Gly Leu Ser Gln Asp Ser

Gly Leu Thr Ile Gly Lys Ser Thr Leu Asn Asn Asp Gly Leu Thr Val

Lys Asp Thr Asn Glu Gln Ile Gln Val Gly Ala Asn Gly Ile Lys Phe

Thr Asn Val Asn Gly Ser Asn Pro Gly Thr Gly Ile Ala Asn Thr Ala

Arg Ile Thr Arg Asp Lys Ile Gly Phe Ala Gly Ser Asp Gly Ala Val

Asp Thr Asn Lys Pro Tyr Leu Asp Gln Asp Lys Leu Gln Val Gly Asn

Val Lys Ile Thr Asn Thr Gly Ile Asn Ala Gly Gly Lys Ala Ile Thr

Gly Leu Ser Pro Thr Leu Pro Ser Ile Ala Asp Gln Ser Ser Arg Asn

Ile Glu Leu Gly Asn Thr Ile Gln Asp Lys Asp Lys Ser Asn Ala Ala

Ser Ile Asn Asp Ile Leu Asn Thr Gly Phe Asn Leu Lys Asn Asn Asn

Asn Pro Ile Asp Phe Val Ser Thr Tyr Asp Ile Val Asp Phe Ala Asn

Gly Asn Ala Thr Thr Ala Thr Val Thr His Asp Thr Ala Asn Lys Thr

Ser Lys Val Val Tyr Asp Val Asn Val Asp Asp Thr Thr Ile His Leu

Thr Gly Thr Asp Asp Asn Lys Lys Leu Gly Val Lys Thr Thr Lys Leu

Asn Lys Thr Ser Ala Asn Gly Asn Thr Ala Thr Asn Phe Asn Val Asn

Ser Ser Asp Glu Asp Ala Leu Val Asn Ala Lys Asp Ile Ala Glu Asn

Leu Asn Thr Leu Ala Lys Glu Ile His Thr Thr Lys Gly Thr Ala Asp

Thr Ala Leu Gln Thr Phe Thr Val Lys Lys Val Asp Glu Asn Asn Asn

Ala Asp Asp Ala Asn Ala Ile Thr Val Gly Gln Lys Asn Ala Asn Asn

Gln Val Asn Thr Leu Thr Leu Lys Gly Glu Asn Gly Leu Asn Ile Lys

Thr Asp Lys Asn Gly Thr Val Thr Phe Gly Ile Asn Thr Thr Ser Gly

Leu Lys Ala Gly Lys Ser Thr Leu Asn Asp Gly Gly Leu Ser Ile Lys

Asn Pro Thr Gly Ser Glu Gln Ile Gln Val Gly Ala Asp Gly Val Lys

Phe Ala Lys Val Asn Asn Asn Gly Val Val Gly Ala Gly Ile Asp Gly

Thr Thr Arg Ile Thr Arg Asp Glu Ile Gly Phe Thr Gly Thr Asn Gly

Ser Leu Asp Lys Ser Lys Pro His Leu Ser Lys Asp Gly Ile Asn Ala

Gly Gly Lys Lys Ile Thr Asn Ile Gln Ser Gly Glu Ile Ala Gln Asn

Ser His Asp Ala Val Thr Gly Gly Lys Ile Tyr Asp Leu Lys Thr Glu

Leu Glu Asn Lys Ile Ser Ser Thr Ala Lys Thr Ala Gln Asn Ser Leu

His Glu Phe Ser Val Ala Asp Glu Gln Gly Asn Asn Phe Thr Val Ser

Asn Pro Tyr Ser Ser Tyr Asp Thr Ser Lys Thr Ser Asp Val Ile Thr

Phe Ala Gly Glu Asn Gly Ile Thr Thr Lys Val Asn Lys Gly Val Val

Arg Val Gly Ile Asp Gln Thr Lys Gly Leu Thr Thr Pro Lys Leu Thr

Val Gly Asn Asn Asn Gly Lys Gly Ile Val Ile Asp Ser Gln Asn Gly

Gln Asn Thr Ile Thr Gly Leu Ser Asn Thr Leu Ala Asn Val Thr Asn

Asp Lys Gly Ser Val Arg Thr Thr Glu Gln Gly Asn Ile Ile Lys Asp

Glu Asp Lys Thr Arg Ala Ala Ser Ile Val Asp Val Leu Ser Ala Gly

Phe Asn Leu Gln Gly Asn Gly Glu Ala Val Asp Phe Val Ser Thr Tyr

Asp Thr Val Asn Phe Ala Asp Gly Asn Ala Thr Thr Ala Lys Val Thr

Tyr Asp Asp Thr Ser Lys Thr Ser Lys Val Val Tyr Asp Val Asn Val

Asp Asp Thr Thr Ile Glu Val Lys Asp Lys Lys Leu Gly Val Lys Thr

Thr Thr Leu Thr Ser Thr Gly Thr Gly Ala Asn Lys Phe Ala Leu Ser

Asn Gln Ala Thr Gly Asp Ala Leu Val Lys Ala Ser Asp Ile Val Ala

His Leu Asn Thr Leu Ser Gly Asp Ile Gln Thr Ala Lys Gly Ala Ser

Gln Ala Asn Asn Ser Ala Gly Tyr Val Asp Ala Asp Gly Asn Lys Val

Ile Tyr Asp Ser Thr Asp Asn Lys Tyr Tyr Gln Ala Lys Asn Asp Gly

Thr Val Asp Lys Thr Lys Glu Val Ala Lys Asp Lys Leu Val Ala Gln

Ala Gln Thr Pro Asp Gly Thr Leu Ala Gln Met Asn Val Lys Ser Val

Ile Asn Lys Glu Gln Val Asn Asp Ala Asn Lys Lys Gln Gly Ile Asn

Glu Asp Asn Ala Phe Val Lys Gly Leu Glu Lys Ala Ala Ser Asp Asn

Lys Thr Lys Asn Ala Ala Val Thr Val Gly Asp Leu Asn Ala Val Ala

Gln Thr Pro Leu Thr Phe Ala Gly Asp Thr Gly Thr Thr Ala Lys Lys

Leu Gly Glu Thr Leu Thr Ile Lys Gly Gly Gln Thr Asp Thr Asn Lys

Leu Thr Asp Asn Asn Ile Gly Val Val Ala Gly Thr Asp Gly Phe Thr

Val Lys Leu Ala Lys Asp Leu Thr Asn Leu Asn Ser Val Asn Ala Gly

Gly Thr Lys Ile Asp Asp Lys Gly Val Ser Phe Val Asp Ser Ser Gly

Gln Ala Lys Ala Asn Thr Pro Val Leu Ser Ala Asn Gly Leu Asp Leu

Gly Gly Lys Val Ile Ser Asn Val Gly Lys Gly Thr Lys Asp Thr Asp

Ala Ala Asn Val Gln Gln Leu Asn Glu Val Arg Asn Leu Leu Gly Leu

Gly Asn Ala Gly Asn Asp Asn Ala Asp Gly Asn Gln Val Asn Ile Ala

Asp Ile Lys Lys Asp Pro Asn Ser Gly Ser Ser Ser Asn Arg Thr Val

Ile Lys Ala Gly Thr Val Leu Gly Gly Lys Gly Asn Asn Asp Thr Glu

Lys Leu Ala Thr Gly Gly Ile Gln Val Gly Val Asp Lys Asp Gly Asn

Ala Asn Gly Asp Leu Ser Asn Val Trp Val Lys Thr Gln Lys Asp Gly

Ser Lys Lys Ala Leu Leu Ala Thr Tyr Asn Ala Ala Gly Gln Thr Asn

Tyr Leu Thr Asn Asn Pro Ala Glu Ala Ile Asp Arg Ile Asn Glu Gln

Gly Ile Arg Phe Phe His Val Asn Asp Gly Asn Gln Glu Pro Val Val

Gln Gly Arg Asn Gly Ile Asp Ser Ser Ala Ser Gly Lys His Ser Val

Ala Ile Gly Phe Gln Ala Lys Ala Asp Gly Glu Ala Ala Val Ala Ile

Gly Arg Gln Thr Gln Ala Gly Asn Gln Ser Ile Ala Ile Gly Asp Asn

Ala Gln Ala Thr Gly Asp Gln Ser Ile Ala Ile Gly Thr Gly Asn Val

Val Ala Gly Lys His Ser Gly Ala Ile Gly Asp Pro Ser Thr Val Lys

Ala Asp Asn Ser Tyr Ser Val Gly Asn Asn Asn Gln Phe Thr Asp Ala

Thr Gln Thr Asp Val Phe Gly Val Gly Asn Asn Ile Thr Val Thr Glu

Ser Asn Ser Val Ala Leu Gly Ser Asn Ser Ala Ile Ser Ala Gly Thr

His Ala Gly Thr Gln Ala Lys Lys Ser Asp Gly Thr Ala Gly Thr Thr

Thr Thr Ala Gly Ala Thr Gly Thr Val Lys Gly Phe Ala Gly Gln Thr

Ala Val Gly Ala Val Ser Val Gly Ala Ser Gly Ala Glu Arg Arg Ile

Gln Asn Val Ala Ala Gly Glu Val Ser Ala Thr Ser Thr Asp Ala Val

Asn Gly Ser Gln Leu Tyr Lys Ala Thr Gln Ser Ile Ala Asn Ala Thr

Asn Glu Leu Asp His Arg Ile His Gln Asn Glu Asn Lys Ala Asn Ala

Gly Ile Ser Ser Ala Met Ala Met Ala Ser Met Pro Gln Ala Tyr Ile

Pro Gly Arg Ser Met Val Thr Gly Gly Ile Ala Thr His Asn Gly Gln

Gly Ala Val Ala Val Gly Leu Ser Lys Leu Ser Asp Asn Gly Gln Trp

Val Phe Lys Ile Asn Gly Ser Ala Asp Thr Gln Gly His Val Gly Ala

Ala Val Gly Ala Gly Phe His Phe

MIGATLSGSAYAQKKDTKHI
AIGEQNQPRRSGTAKADGDR
AIAIGENANAQGGQAIAIGS
SNKTVNGSSLDKIGTDATGQ
ESIAIGGDVKASGDASIAIG
SDDLHLLDQHGNPKHPKGTL
INDLINGHAVLKEIRSSKDN
DVKYRRTTASGHASTAVGAM
SYAQGHFSNAFGTRATAKSA
YSLAVGLAATAEGQSTIAIG
SDATSSSLGAIALGAGTRAQ
LQGSIALGQGSVVTQSDNNS
RPAYTPNTQALDPKFQATNN
TKAGPLSIGSNSIKRKIINV
GAGVNKTDAVNVAQLEAVVK
WAKERRITFQGDDNSTDVKI
GLDNTLTIKGGAETNALTDN
NIGVVKEADNSGLKVKLAKT
LNNLTEVNTTTLNATTTVKV
GSSSSTTAELLSDSLTFTQP
NTGSQSTSKTVYGVNGVKFT
NNAETTAAIGTTRITRDKIG
FARDGDVDEKQAPYLDKKQL
KVGSVAITIDNGIDAGNKKI
SNLAKGSSANDAVTIEQLKA
AKPTLNAGAGISVTPTEISV
DAKSGNVTAPTYNIGVKTTE
LNSDGTSDKFSVKGSGTNNS
LVTAEHLASYLNEVNRTADS
ALQSFTVKEEDDDDANAITV

AKDTTKNAGAVSILKLKGKN
GLTVATKKDGTVTFGLSQDS
GLTIGKSTLNNDGLTVKDTN
EQIQVGANGIKFTNVNGSNP
GTGIANTARITRDKIGFAGS
DGAVDTNKPYLDQDKLQVGN
VKITNTGINAGGKAITGLSP
TLPSIADQSSRNIELGNTIQ
DKDKSNAASINDILNTGFNL
KNNNNPIDFVSTYDIVDFAN
GNATTATVTHDTANKTSKVV
YDVNVDDTTIHLTGTDDNKK
LGVKTTKLNKTSANGNTATN
FNVNSSDEDALVNAKDIAEN
LNTLAKEIHTTKGTADTALQ
TFTVKKVDENNNADDANAIT
VGQKNANNQVNTLTLKGENG
LNIKTDKNGTVTFGINTTSG
LKAGKSTLNDGGLSIKNPTG
SEQIQVGADGVKFAKVNNNG
VVGAGIDGTTRITRDEIGFT
GTNGSLDKSKPHLSKDGINA
GGKKITNIQSGEIAQNSHDA
VTGGKIYDLKTELENKISST
AKTAQNSLHEFSVADEQGNN
FTVSNPYSSYDTSKTSDVIT
FAGENGITTKVNKGVVRVGI
DQTKGLTTPKLTVGNNNGKG
IVIDSQNGQNTITGLSNTLA
NVTNDKGSVRTTEQGNIIKD
EDKTRAASIVDVLSAGFNLQ
GNGEAVDFVSTYDTVNFADG
NATTAKVTYDDTSKTSKVVY
DVNVDDTTIEVKDKKLGVKT
TTLTSTGTGANKFALSNQAT
GDALVKASDIVAHLNTLSGD
IQTAKGASQANNSAGYVDAD
GNKVIYDSTDNKYYQAKNDG
TVDKTKEVAKDKLVAQAQTP
DGTLAQMNVKSVINKEQVND
ANKKQGINEDNAFVKGLEKA
ASDNKTKNAAVTVGDLNAVA
QTPLTFAGDTGTTAKKLGET
LTIKGGQTDTNKLTDNNIGV
VAGTDGFTVKLAKDLTNLNS
VNAGGTKIDDKGVSFVDSSG
QAKANTPVLSANGLDLGGKV
ISNVGKGTKDTDAANVQQLN
EVRNLLGLGNAGNDNADGNQ
VNIADIKKDPNSGSSSNRTV
IKAGTVLGGKGNNDTEKLAT
GGIQVGVDKDGNANGDLSNV
WVKTQKDGSKKALLATYNAA
GQTNYLTNNPAEAIDRINEQ
GIRFFHVNDGNQEPVVQGRN
GIDSSASGKHSVAIGFQAKA
DGEAAVAIGRQTQAGNQSIA

IGDNAQATGDQSIAIGTGNV
VAGKHSGAIGDPSTVKADNS
YSVGNNNQFTDATQTDVFGV
GNNITVTESNSVALGSNSAI
SAGTHAGTQAKKSDGTAGTT
TTAGATGTVKGFAGQTAVGA
VSVGASGAERRIQNVAAGEV
SATSTDAVNGSQLYKATQSI
ANATNELDHRIHQNENKANA
GISSAMAMASMPQAYIPGRS
MVTGGIATHNGQGAVAVGLS
KLSDNGQWVFKINGSADTQG
HVGAAVGAGFHF